

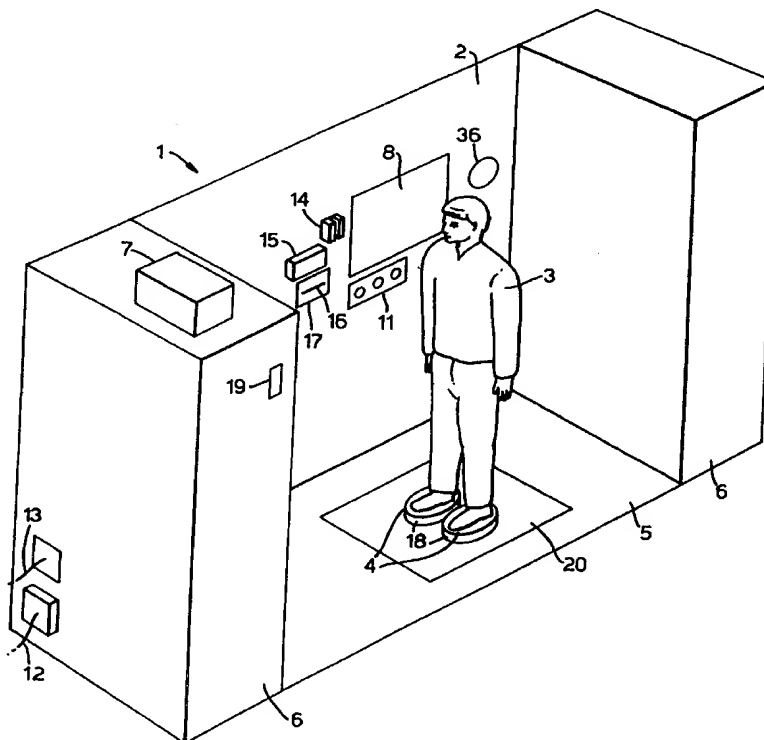


## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>H04N 1/00</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 98/28908</b> <b>(43) International Publication Date:</b> 2 July 1998 (02.07.98)
<b>(21) International Application Number:</b> PCT/GB97/03515 <b>(22) International Filing Date:</b> 22 December 1997 (22.12.97) <b>(30) Priority Data:</b> 9626825.5      24 December 1996 (24.12.96)      GB <b>(71)(72) Applicant and Inventor:</b> CRAMPTON, Stephen, James [GB/GB]; 9 Broadfields, Goffs Oak, Waltham Cross, Herts EN7 5JU (GB).		<b>(81) Designated States:</b> AT, AU, BG, BR, CA, CH, CN, CU, CZ, DE, DK, ES, FI, GB, HU, IL, JP, KP, KR, MX, NO, NZ, PL, PT, RU, SE, SG, TR, UA, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  <b>Published</b> <i>With international search report.          Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>

**(54) Title:** AVATAR KIOSK**(57) Abstract**

The device (100) is rugged, self-contained and compact with an internal area (102) for a person (103) to stand in. Markings (104) for foot positioning are provided on the floor (105) of the internal area. Sensing equipment (106) enclosed in the kiosk is situated at one end. The sensing equipment is connected to a computer (107). A display monitor (108) is connected to the computer. The computer is connected to the operator panel (111). The computer is connected to a telecommunications line (112). The device (100) is connected to mains electricity (113). A credit card swipe reader (114) is connected to the computer. A colour printer (115) connected to the computer outputs a colour image of the Avatar (116) into a tray (117) where the person can recover it. Electronic weighing scales (118) are situated under the floor (105) and are connected to the computer. A microphone (119) is connected to the computer. A loudspeaker (120) is connected to the computer. Internal lighting (121) is connected to the electricity supply. The exterior of the kiosk comprises backlit advertising panels (122).



***FOR THE PURPOSES OF INFORMATION ONLY***

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

<b>AL</b>	Albania	<b>ES</b>	Spain	<b>LS</b>	Lesotho	<b>SI</b>	Slovenia
<b>AM</b>	Armenia	<b>FI</b>	Finland	<b>LT</b>	Lithuania	<b>SK</b>	Slovakia
<b>AT</b>	Austria	<b>FR</b>	France	<b>LU</b>	Luxembourg	<b>SN</b>	Senegal
<b>AU</b>	Australia	<b>GA</b>	Gabon	<b>LV</b>	Latvia	<b>SZ</b>	Swaziland
<b>AZ</b>	Azerbaijan	<b>GB</b>	United Kingdom	<b>MC</b>	Monaco	<b>TD</b>	Chad
<b>BA</b>	Bosnia and Herzegovina	<b>GE</b>	Georgia	<b>MD</b>	Republic of Moldova	<b>TG</b>	Togo
<b>BB</b>	Barbados	<b>GH</b>	Ghana	<b>MG</b>	Madagascar	<b>TJ</b>	Tajikistan
<b>BE</b>	Belgium	<b>GN</b>	Guinea	<b>MK</b>	The former Yugoslav Republic of Macedonia	<b>TM</b>	Turkmenistan
<b>BF</b>	Burkina Faso	<b>GR</b>	Greece			<b>TR</b>	Turkey
<b>BG</b>	Bulgaria	<b>HU</b>	Hungary	<b>ML</b>	Mali	<b>TT</b>	Trinidad and Tobago
<b>BJ</b>	Benin	<b>IE</b>	Ireland	<b>MN</b>	Mongolia	<b>UA</b>	Ukraine
<b>BR</b>	Brazil	<b>IL</b>	Israel	<b>MR</b>	Mauritania	<b>UG</b>	Uganda
<b>BY</b>	Belarus	<b>IS</b>	Iceland	<b>MW</b>	Malawi	<b>US</b>	United States of America
<b>CA</b>	Canada	<b>IT</b>	Italy	<b>MX</b>	Mexico	<b>UZ</b>	Uzbekistan
<b>CF</b>	Central African Republic	<b>JP</b>	Japan	<b>NE</b>	Niger	<b>VN</b>	Viet Nam
<b>CG</b>	Congo	<b>KE</b>	Kenya	<b>NL</b>	Netherlands	<b>YU</b>	Yugoslavia
<b>CH</b>	Switzerland	<b>KG</b>	Kyrgyzstan	<b>NO</b>	Norway	<b>ZW</b>	Zimbabwe
<b>CI</b>	Côte d'Ivoire	<b>KP</b>	Democratic People's Republic of Korea	<b>NZ</b>	New Zealand		
<b>CM</b>	Cameroon			<b>PL</b>	Poland		
<b>CN</b>	China	<b>KR</b>	Republic of Korea	<b>PT</b>	Portugal		
<b>CU</b>	Cuba	<b>KZ</b>	Kazakstan	<b>RO</b>	Romania		
<b>CZ</b>	Czech Republic	<b>LC</b>	Saint Lucia	<b>RU</b>	Russian Federation		
<b>DE</b>	Germany	<b>LI</b>	Liechtenstein	<b>SD</b>	Sudan		
<b>DK</b>	Denmark	<b>LK</b>	Sri Lanka	<b>SE</b>	Sweden		
<b>EE</b>	Estonia	<b>LR</b>	Liberia	<b>SG</b>	Singapore		

## AVATAR KIOSK

This invention relates to a kiosk apparatus and method for capturing data on the surface of a person and creating an Avatar of that person.

Several person scanning machines have been built to scan the surface of a person. Some person scanning machines are limited in scope to parts of a person such as the head or hand.

The amount of time to scan a person varies in different machines. Some person scanning machines capture all the data quickly enough to freeze any movement. Other machines take several seconds.

Since parts of the human body obscure other parts of the body, it is necessary to either have multiple sensors or relative movement between the sensor and the person. The main methods of movement that have been used in person scanning machines are cylindrical rotation of the person or linear translation of several sensors.

All machines allow the person to adopt any posture during scanning. However most person scanning machines have a restricted data capture volume and any parts of the body outside that volume are missed.

All person scanning machines capture a representation of the 3D shape of the body. Some machines capture other information such as colour as well.

Most of these person scanning machines are very large in size and often occupy a dedicated room. None of these machines are self-contained in a small kiosk.

Passport photo kiosks are common in public places. They are usually self-contained with an access for a person to enter.

All of these person scanning machines have been designed for a research or laboratory environment. They are usually unrobust and not designed for dirty environments in which they can be knocked and subject to high degrees of wear and tear.

Since most person scanning machines use sensors that are a long distance away from the subject and allow the subject to adopt any posture, it is usual for data in many areas to be obscured.

All existing person scanning machines require highly trained personnel to operate them. None are intended for use by a passer-by in the same way that a passport photo kiosk is operated.

None of the systems include electronic weighing means. Weight cannot be accurately estimated from the known 3D shape of a person, because different people have different amount of fat, bone and tissue, each of which have different densities.

The raw output of all these person scanning machines is a large number of 3D points and texture information in different forms. Special software is then used to convert the raw data into the required forms. Such special software needs a skilled operator to use it.

3D Scanners Ltd in the UK developed a whole body scanner called PERSONA in 1993. The person being scanned stood on a rotating table. Two sensors projected laser stripes through the axis of rotation of the table. As the person was turned, the shape of the body was captured using the principle of laser stripe triangulation. Considerable amounts of surface are obscured using this design.

Cyberware Labs Inc in the USA developed a whole body scanner in 1994. The person remained stationary whilst four sensors which projected horizontal laser stripes descended down gantries. The colour of the person was also recorded. The process takes around 17 seconds. This is a long time for a person to maintain the same position. Also, data is not captured for instance the top of the head and the tops of the shoulders are missed.

3D Scanners Ltd in the UK developed a scanner called ModelMaker in 1996. ModelMaker captures 3D information using a laser stripe sensor and captures texture information using a colour camera. The ModelMaker software can automatically create a 3D polygonal model from the data and texture it with the colour images. The automatic algorithm pipeline for doing that first creates triangular meshes from each sweep of the laser stripe sensor. Then it forms an implicit surface from the triangular meshes. Then it generates a single polygonal mesh from the implicit surface. Holes in the mesh are automatically filled. The mesh is smoothed and polygon reduced. The colour images are mapped onto the mesh to produce texture maps.

Thomas Gentils and Adrian Hilton at Surrey University in Guildford produced a report placed in the public domain in November 1997 on automatically generating an Avatar from four colour images of a person and a standard 3D polygonal model. Their method does not use any 3D measurement. There is a single colour camera. The person makes a similar pose in each of four

orthogonal orientations for four images. The background behind the person is a blue screen. They use standard image processing algorithms for extracting silhouettes and landmarks from the images based on knowledge of the background being blue. They extract silhouettes and similar landmarks from a standard Avatar 3D model. They use standard image processing algorithms for 2D to 2D mapping the silhouettes and landmarks of the standard Avatar 3D model to the silhouettes and landmarks from the images. They use standard image processing algorithms for using the mapped silhouettes and landmarks to deform the standard Avatar 3D model into an untextured Avatar 3D model of the user. They also use standard image processing algorithms and the mapped silhouettes and landmarks to map the image textures onto the untextured Avatar 3D model of the user to produce a textured Avatar 3D model of the user. This works well unless the person is wearing blue clothing such as blue jeans. If blue clothing is worn the silhouette is indistinguishable from the clothing and a faulty Avatar is generated.

Several companies have developed 3D cameras which take a large number of measurements in a short period of time and have used several 3D cameras at the same time to scan a person. 3D cameras have never found commercial success, mainly due to technical limitations.

There are several systems for capturing facial expressions. They mostly use ordinary cameras and landmarks stuck on key points on a person's face for use in real-time in which a virtual actor driven by a powerful computer mimics the person whose facial expressions are being recorded.

Avatars, also known as virtual humans, are used to represent a person in a virtual environment. Many different 'anthropomorphic' software models of Avatars have been developed.

There are three main types of Avatar representations: sprites, polygons and high-level surfaces. Sprites are a pseudo-3D representation in which an Avatar consists of several 2D images taken from different angles. The nearest image to the direction of observation is used to represent an Avatar in a virtual environment. A human sprite Avatar can be created from a number of photos of that person. Sprites are mainly used in games where there is no close-up of the character being represented. Most virtual environments are 3D and have close-up navigation. The sprite representation in a virtual environment is very poor when the viewpoint into the environment is close to the position of the sprite image. Polygons are commonly used and can include texture maps. High-level surface representations are not used very often.

An international standard for an Avatar model has been drafted by the VRML Consortium's 'Humanoid Animation Working Group'. This draft "Specification for a Standard VRML Humanoid, Version 1.0" was published in September 1997 and works under current VRML 2.0 compliant browsers.

It is very difficult to manually create polygon and high-level surface Avatars similar to real people. Using photos and a mouse, it can take many days of skilled work to create a poor representation of that person.

Most Avatars in use in virtual worlds today bear no resemblance to the person they represent. The Avatars in use that look closest to the people they represent are sprites made from photos and standard polygonal models on which a bit-mapped image of the person's face is wrapped.

In the real world, a person's identity is guaranteed by a card or passport which usually bears a photo of the person. Photos have been mandatory in passports in some European countries for over 80 years. Financial guarantees are often given in the form of credit cards with encoded information on them which are checked for validity against central records. On the Internet, a high-security encryption system is offered by many software vendors to secure information transmissions.

According to the present invention, there is provided a novel kiosk in which the Avatar of a person is automatically generated. The kiosk [1] is free-standing and has a central area [2] for a person [3] to stand in. Foot positioning guidance means [4] are provided on the floor [5] of the central area. Sensing equipment [6] is situated in front of and behind the person. The sensing equipment is connected to computing means [7]. A display [8] is connected to the computing means [9]. The computing means [9] is connected to an operator input means [11]. The computing means [9] is connected to a telecommunications means [12]. The kiosk [1] is connected to a source of power [13]. A means for reading encoded information [14] is connected to the computing means [9]. A printer [15] may be connected to the computing means [9] which outputs a printed document [16] into a tray [17]. Electronic weight measurement means [18] may be positioned under the foot guidance means [4] and be connected to the computing means [9]. A sound recording and digitising means [19] may be connected to the computing means [9]. A rolling walkway [20] could be installed in the floor [5] of the central area.

The person stands in a specified posture [21] with his feet [22] positioned using the foot positioning guidance means [4] so that the person faces in the specified direction with his feet in the specified places. The specified posture may require that the arms [23] are held away from the body to either side such that the arms

and the main body [24] are roughly in the same, approximately vertical plane P. The hands [25] may be specified to have the fingers [26] spread and the palms facing backwards (thumbs [27] pointing down) such that the hands and fingers are roughly in the same plane P. The legs [33] and the back [34] may be specified as being straight. A head positioning guidance means [28] may be used to locate the head [29] such that it is roughly vertical and not too far forward or back.

An example of the foot positioning guidance means [4] is two footmark graphics on the floor [5] of the kiosk [1]. Such graphics give both position and direction in a clearly recognisable form. A second example is two, slightly raised foot-shaped areas. Such areas raise the feet slightly and enable them to be better scanned.

An example of such head positioning guidance means is a beam of light [30] projected vertically downwards such that in the optimum posture, the person will see the incidence of the beam of light [31] with the tip of his nose [32] in a mirror [37].

An example illustration of the specified posture [21] which may comprise written words or graphics or both may be provided as an aid to the person to help him attain the specified posture. Alternatively or additionally, a sound projection means [36] may be connected to the computing means [9] to issue posture instructions which may either be pre-recorded or issued in a computer-generated voice.

The user input means [11] may for example be a touch screen surface on the display [8] or a keyboard or a panel with a number of buttons or any combination of these or other user interface means. It is used to select any options the kiosk offers. The user may enter the electronic mail address to which the Avatar dataset should be sent or a password for identifying a retrieval from a central server. More extensive identity information such as personal and financial information may be entered.

The telecommunications means [12] is used to deliver the Avatar. It or other means is used to carry out the credit card transaction. It can also automatically notify a central service computer of any malfunction; similarly any lack of response to an 'are you there kiosk' enquiry from the central service computer can be checked by means of a visit.

Alternatively, an Avatar may be delivered to the user on a computer readable medium such as a diskette or writable CD-ROM through a slot into a tray. Users may want to take away the Avatar on a computer readable medium rather than have it delivered on a network.

The means for reading encoded information [14] may be a credit card 'swipe' reader. The use of credit cards has the advantage of enabling high amounts to be charged for Avatar creation without making the kiosk [1] an attractive target for thieves which would be the case if coins or banknotes were inserted. A second advantage of credit cards is that the kiosk would not require regular visits to empty it of money. If electronic money becomes generally accepted then the kiosk would support that and other forms of credit card or cash would not be used.

The display [8] may be a computer monitor. It may be a large LCD flatscreen or any other form of pixel display technology. It is used for a variety of functions including:

- offering options for the user to select
- acknowledging selections / payments
- echoing user input
- confirming user electronic mail address availability
- issuing instructions
- displaying the resulting Avatar
- displaying an address/password where the Avatar dataset may be downloaded from
- advertising what the kiosk does pre-sale
- advertising other products post-sale

A printer [15] may provide the user with a take-away image of his Avatar dataset or a password and address for downloading his Avatar from a central server.

The electronic weight measurement means [18] can measure the weight of the person and his clothes.

Sound recording and digitising means [19] may be used to capture the voice of the person as he makes sounds. A combination of speaking specified words and making known sounds such as laughing might enable the main characteristics of a person's voice to be determined such that he can be more accurately represented to others in virtual environments ie his Avatar sounds like him.

The rolling walkway [20] could be used to capture motion such as walking. The walkway could be motor driven or inactive. It could be uni-directional or bi-directional. A bi-directional walkway would enable more complex motions such as turning to be captured. The sensing equipment [6] would need to be able to identify and follow the motion of the person. Kiosks with motion capture capabilities would be larger than those without. If the kiosk cannot capture motion, then using the user input means [11], the person can specify the characteristic way



in which he moves by answering questions and selecting a walk similar to his from a small library of electronic video displayed on the display [8]. It is difficult to know your own walk, so the presence of a friend may help. The question-answering and video selection process may be interactive with dynamic animations of the person's Avatar being shown after each question is answered.

The sensing equipment [6] can capture the person's shape and colour over a period of time. It can also capture the person's face with several different expressions such as happy, sad, angry and laughing in several sequential sessions. Capturing expressions provides more information such that the reality of the person's Avatar may be enhanced. Application software can use the Avatar datasets including the person's expressions to visually recreate moods.

The person may be offered the opportunity to view an Avatar that has been generated of that person on the display and accept it or to have new data captured and a new Avatar generated. This opportunity may be repeated one or more times.

The sensing equipment [6] can be of many different types. It can consist of a number of static 3D cameras. The 3D cameras could use fringe or other types of structured light. It could consist of a number of laser stripe sensors translated on two linear axes. Colour cameras and flashlights could be used to capture the colours of the person.

The sensing equipment [6] needs to be registered so that the data from all the sensors can automatically be registered in one coordinate system. This would be done by means of a simple alignment procedure after installation. The best method is probably the capturing of a known artifact. It assumes that all the sensors are calibrated and in two known frames. One such artifact [40] consists of 8 spheres [41] in a known relation to each other. The 8 spheres are connected by four horizontal tubes [42] and a series of vertical wires [43]. A weight [44] on rubber string [45] rests on the floor to stop the artifact swinging. This artifact [40] can be hung from the top of the kiosk [46] and also has the capability of being stored in a relatively small travelling case.

One embodiment of the sensing equipment [6] is two linear axes [50a, 50b] on either side of the person. The main data capture method is laser stripe. Laser stripe is low-cost, relatively fast and does not give ambiguous results. The sensing equipment mounted on a linear axis travels just over a quarter of the person's height. There are eight sensors [57] on each axis. On each axis, sensors may be mounted in pairs to give better all round coverage, reduce stripe length and reduce occlusion. In this way the capture time is quartered, the time between scans for the sensing equipment to retract is quartered and a capture time of less than 5

seconds is achievable. This leads to a total of 16 sensors. Each sensor has one stripe [51] pointing at a small angle (5 to 10 degrees) upwards and a second stripe [52] pointing at a small angle (5 to 10 degrees) downwards. The stripes are switched so only one stripe is illuminated at a time thus removing the chance of ambiguous data. Some sensors have two cameras [53, 54], other sensors have one camera [either 53 or 54]. Each stripe may be generated by a laser diode collimator [55] with a visible red 670 nm laser diode and a fixed cylindrical optic [58]. Each laser diode may be Class II which is eye-safe. The specification of a known, compact posture enables the sensing equipment to be made much smaller.

In this embodiment, 16 colour cameras are used with 16 flash lights to record the colour of the person. The colour images are taken at three points: at the top of the travel, in the middle and at the bottom. At each point, the eight colour cameras and flashes on one side operate synchronously and shortly afterwards the eight colour cameras and flashes on the other side operate synchronously. This avoids cameras on one side being blinded by the flashes on the other side. The 16 colour cameras [58] and flashes [59] may be mounted on the 16 sensors [57].

In this embodiment, the two axes are motor driven by either servo or stepper motors. The ballscrew or other drive mechanism need only be just over a quarter of a person's height. A single linear bearing is used on each axis to maintain alignment, straightness and reduce cost.

The large quantity of 3D points resulting from the scan and the colour data must be automatically converted into an Avatar. Most sensors also provide some estimate of surface normal for most 3D data points.

A closed, texture mapped, 3D polygonal mesh Avatar can be automatically constructed from the data. One automatic algorithm pipeline that can be used is this:

- for each sensor create a 2.5D triangular sub-mesh from the data for that sensor (these sub-meshes will often have holes in them)
- populate a multi-resolution, volumetric data structure with the 2.5D triangular sub-meshes to form an implicit surface
- use the implicit surface to generate a single polygonal mesh
- fill any holes in the mesh
- decimate (polygon reduce) the polygonal mesh to reduce the quantity of data to the desired amount
- smooth the mesh using a median smoothing algorithm to maintain features
- reverse render the colour images onto the mesh to produce texture maps

This polygon Avatar is automatically generated and may be used at different levels of detail in a variety of virtual environments. The random nature of the polygons make it difficult to use the Avatar for animation of body and face.

To animate an Avatar's body, the identification of joints and the construction of a jointed model is required. One automatic algorithm pipeline that can be used to achieve an animatable Avatar is this:

- find landmarks from 3D feature detection using heuristic rules based on the known posture. Landmarks might include head, nose, ears, feet, hands, fingers
- fit a 'standard' human jointed model between the landmarks
- optimise the jointed model to the polygon model using a least energy approach

An Avatar's face can be animated based on a standard model or if several expressions of the person have been captured a special model can be generated. The identification of facial features is required. One automatic algorithm pipeline that can be used to achieve facial animation is this:

- find landmarks on face using colour textures eg eyebrow ends, mouth corners, nose, eyes, chin
- work out their position relative to each other in the different expressions
- normalise the different expressions in space ie register them together

In another embodiment of this invention, a number of cameras [63] are used to capture images of a person [3]. To capture the person in one pose, a large kiosk [60] is required. This is because the cameras must be placed on at least 4 sides of the person. Such a large kiosk [60] has high manufacture, transport, installation and space rental costs as well as a lower degree of reliability. A kiosk [61] has cameras on 2 sides and requires only 2 poses; it is much smaller. A kiosk [62] has cameras on 1 side only and requires 4 poses; it has the lowest manufacture, transport, installation and space rental costs as well as a higher degree of reliability.

To further reduce the size of the kiosk, the optical path [64] of a camera [63] may be folded through an angle by means of a mirror [65].

A further reduction in size of the kiosk is achieved by using a set of cameras on each side arranged in a grid. There is an additional benefit in that this increases the total image information if cameras of the same resolution are used. However any quality benefits may be reduced by the increases in cost and reliability. It is possible that for a given required amount of image data from each side that using

one, expensive high resolution camera may be better than several, cheaper low resolution cameras. The camera to person distance  $d$ [number of cameras] is reduced with an increasing number of cameras [63] producing images [66] such that  $d[1] > d[2] > d[4]$ . The cameras may be individually calibrated and the set of cameras aligned to each other using techniques known to those skilled in the art. Such a set of calibrated and aligned cameras significantly aid the process of combining the neighbouring and overlapping images of one side of the person such that there is higher quality of silhouette extraction and texture mapping in the Avatar generation process. This invention is not limited to these three arrangements of cameras and is applicable to any arrangements of cameras. If several sides are used, then several sets of cameras will be required.

It is also a purpose of this invention to overcome the problem of the wearing of blue clothes in front of a blue background. A background [72] is used with retroreflective material. This background material [72] may be manufactured by densely sewing or weaving retroreflectors onto cloth. The lens [69] of the camera [63] is surrounded by light generators such as light emitting diodes (LEDs) of two significantly different wavelengths [70,71]. An image of the person [3] is taken with camera [63] and background [72] with LEDs [70] illuminated and LEDs [71] not illuminated. A short time later such that the person will not have been able to move significantly, a second image is taken with the same camera but with LEDs [71] illuminated and LEDs [70] not illuminated. For both images there is also normal white lighting [73] to properly illuminate the person. In the first image, the retroreflective background appears to be the colour of the LEDs [70] and in the second image, the retroreflective background appears to be the colour of the LEDs [71]. The two images of the person are very similar. Simple image processing algorithms can be used to combine the silhouettes reliably using thresholding of the characteristic colour unless an item of clothing has both LED colour characteristics. The scope of this invention is not limited to this single method of overcoming the blue background problem and this invention is applicable to any other method of overcoming this problem such as by taking 3 images with 3 different sets of LEDs and combining them using algorithms that blend rather than threshold.

At present each software application that uses avatars has its own format. Some formats are static, some animatable, some include expressions, some personal and financial information.

An Avatar Kiosk can output all its Avatars in a standard format of its own specification: 'AVSTAN'. The AVSTAN format may use any or all of the following defined data structures and any number of data instances for the structures:

- static shape topology

- texture maps registered to the shape topology
- landmarks indexed to the shape topology
- standard joint topology
- positions of joints relative to the shape topology
- expression shape deformations of static shape topology eg angry, happy
- characteristic motions eg walking
- contact information eg e-mail address, geographical locations, telecomms numbers
- financial information, spending habits
- personal information such as weight, age, likes, habits, family
- medical, health information
- curriculum vitae (resume)
- community record eg fines, imprisonments
- relationship information eg psychological profile, attraction profile

By fully defining the AVSTAN format and having it include a wide range of data, different software packages can then quickly generate a representation of a person entering their virtual environment in that software's format from the person's AVSTAN dataset and knowledge of the standard AVSTAN format. The advantage of a standard is that it is recognised in all environments worldwide.

What a person wears is relevant. Bulky clothes obscure a person's true shape. If an Avatar is generated of someone wearing bulky clothes, then it will not be possible to use the Avatar data for sizing in teleshopping or for many health and beauty applications. It will also be technically difficult for the Avatar to change into other virtual clothes. One solution is for a person to have a different Avatar made with each new clothes outfit and hairstyle. The alternative of few clothes requires the removal of most clothing. It is likely that an Avatar Kiosk for applications requiring the output of accurate body shape would need to be situated where people are happy to remove their clothes such as in the changing area of a clothes shop or a sports centre.

It is also a part of this invention to circumvent the reluctance of people to take their clothes off when using a kiosk by means of a novel apparatus and method. Weight measurement provides a measurement of the weight of the person and his clothes. An estimate may be made for the person's actual weight by using a measured metric obtained from clothing retailers for the percentage of weight of clothes that a person wears relative to his body weight in different temperatures. A measured metric of the average density of a person may be obtained from published anthropometric data for different height to weight ratios. The user's 3D Avatar is an enclosed volume that is easily calculated by someone skilled in the art. The volume that it should be for the adjusted weight can be calculated using

the average density for the height to weight ratio. The Avatar can then be deformed to this volume to produce an estimated naked Avatar using standard deformable algorithms which are well known to those skilled in the art. The textures of the estimated naked Avatar can be generated using facial textures to estimate the person's skin colour. This estimated naked Avatar can be used for clothes fit applications. It overcomes the problem of people's reluctance to take their clothes off in a public place. People wanting accurate, estimated naked avatars should use the kiosk whilst wearing minimal and very tight fitting clothing. This extension of the invention makes kiosks for clothes fit viable.

The user can specify whether and how much of his Avatar identity be published in a directory similar to a telephone directory. The user can also specify whether his Avatar identity be made available for direct marketing purposes. If the value of the data is high to direct marketing companies then the cost for using the Avatar Kiosk may be reduced. As the entrypoint to Cyberspace, the Avatar Kiosk is likely to be of high strategic value to many types of company.

The high quality of the Avatars produced in the kiosk will improve the reality of virtual environments and is likely to enhance their uptake by the public in different applications. Like passport photo kiosks the cost is low. Unlike passport photo kiosks a polaroid Avatar cannot be taken at home - the specialised equipment in the Avatar kiosk is not available in homes and offices. It is likely that the Avatar kiosk will have high margins and a high market-share of all those people requiring Avatars.

According to the invention there is also provided a method for generating an Avatar of a person in the kiosk.

The invention will now be described, by way of example only, with reference to the accompanying Figures:

Figure 12 is an outline of the system. The device [100] is rugged, self-contained and compact with an internal area [102] for a person [103] to stand in. The device for purposes of clarity is shown in Figure 12 with the ceiling, entry and one side removed. The entry may be a simple entry gap in one side wall with a curtain across it. Markings [104] for foot positioning are provided on the floor [105] of the internal area. Sensing equipment [106] enclosed in the kiosk is situated at one end. The sensing equipment is connected to a means for programmable processing (a computer) [107]. A display monitor [108] is connected to the computer. The computer is connected to the operator panel [111]. The computer is connected to a telecommunications line [112]. The device [100] is connected to mains electricity [113]. A credit card swipe reader [114] is

connected to the computer. A colour printer [115] connected to the computer outputs a colour image of the Avatar [116] into a tray [117] where the person can recover it. Electronic weighing scales [118] are situated under the floor [105] and are connected to the computer. A microphone [119] is connected to the computer. A loudspeaker [120] is connected to the computer. Internal lighting [121] is connected to the electricity supply. The exterior of the kiosk comprises backlit advertising panels [122].

Figure 2 is an outline of a preferred posture. The person [3] stands on the markings [4] with his back [34] straight. His feet [22] face forward and his legs [33] are straight. His arms [23] are stretched out to either side with his hands [25] with palms facing backwards, thumbs [27] pointing downwards and fingers separated. His head [29] is located with a beam of light [30] from a projector [28] makes a spot of light [31] on his nose [32]. The same posture is used for each of the four images.

Figure 7 shows plan views of 3 possible kiosk layouts. Kiosk [62] is the preferred layout requiring cameras [63] on one side and four postures of the person [3].

Figure 8 is a preferred optical path [64] between the camera [63] and the person [3] which is folded using a mirror [65].

Figure 9 shows views of several camera combinations. The preferred number of cameras would be determined by the estimated cost of the components at the time of manufacture.

Figure 10 shows a ring of LEDs [70,71] around the lens [69] of a camera [63]. The LEDs [70] and LEDs [71] have significantly different wavelengths.

Figure 11 shows a camera [63] with LEDs [70,71] imaging a person [3] with the floor, back and roof of the kiosk covered in a retroreflective background material [72].

Figure 1 is an outline of a system. The device [1] is self-contained with a central area [2] for a person [3] to stand in. Two raised footstands [4] are provided on the floor [5] of the central area. Sensing equipment [6] enclosed in the device [1] is situated in front of and behind the person. The sensing equipment [6] is connected to a computer [7]. A display monitor [8] is connected to the computer. The computer is connected to the operator panel [11]. The computer [7] is connected to a telephone line [12]. The device [1] is connected to mains electricity [13]. A credit card swipe reader [14] is connected to the computer. A colour printer [15] connected to the computer outputs a colour image of the Avatar [16] into a tray [17] where the person can recover it. Electronic scales [18] are situated

under each footstand [4] and are connected to the computer. A microphone [19] is connected to the computer.

Figure 3 is an illustration of the registration artifact [40]. Eight spheres [41] are hung in an accurately known orientation relative to each other from the top [46] of the central area. Four horizontal tubes [42] connect the spheres [41] in pairs. The tubes [42] are connected together by vertical wires [43]. A weight [44] on a rubber string rests on the floor [5] to help stop the artifact swinging.

Figure 4 is an illustration of the two linear axes [50a,50b] with 16 sensors [57] in four groups of four sensors.

Figure 5 is an illustration of a sensor [57]. Each sensor [57] has one stripe [51] pointing at a small angle (5 to 10 degrees) upwards and a second stripe [52] pointing at a small angle (5 to 10 degrees) downwards. The stripes are switched so only one stripe is illuminated at a time thus removing the chance of ambiguous data. All sensors have two cameras [53, 54]. Each stripe is generated by a laser diode collimator [55] with a visible red 670 nm laser diode and a fixed cylindrical optic [58]. Each laser diode is Class II which is eye-safe. A colour camera [58] and a flash light [59] are mounted on the sensor [57].

Figure 6 is a plan view of the device [1] with the person [3] standing in the recommended posture such that his body, head, arms, hands and legs are approximately aligned with a near-vertical plane P. The sensors [57] are rigidly attached to one of the two axes [50a,50b]. The positions of the four sets of sensors [57] in Figure 6 can be optimised to recover the maximum amount of the surface area of the person.

A configuration for the Avatar Kiosk is shown in Figure 12. This configuration is only one of many modes for carrying out the invention.

The invention can be applied in the form of kiosks. Such kiosks can be placed in accessible positions such as airports, amusement centres, shops and health centres throughout the world. The kiosks will be connected to a telecommunications network to assure the transfer of the financial transaction and the Avatar dataset. People can use a credit card or electronic money to pay for the transaction. The whole process will take a few minutes.

Applications of the resulting Avatar datasets include: the 'passport' or person's identity in cyberspace or virtual environments, health and beauty treatment;



teleshopping, virtual meetings, lonely hearts, games, clothing size specifications, medical treatments and guaranteeing financial transactions.

It is likely that people will wish to repeat the process at different times in life or in wearing different clothes or with different hairstyles and makeup.

## CLAIMS

1. A kiosk apparatus for capturing data on the surface of a person to provide an Avatar thereof, the apparatus comprising:

means for programmable processing;  
means for capturing data on the surface of a person;  
means for generating an Avatar that has a resemblance of the person from the captured data;  
means for transmitting the Avatar.

2. A kiosk apparatus according to claim 1, characterised in that the data capture means comprises means for illuminating the person and means for capturing one or more images of the person.

3. A kiosk apparatus according to claim 2, characterised in that the means for capturing one or more images of the person is one or more cameras.

4. A kiosk apparatus according to claim 3, characterised in that one camera is positioned such that the visible surface of one side of a person whose dimensions are the largest dimensions for which the kiosk is designed is captured on a single image.

5. A kiosk apparatus according to claim 3, characterised in that two or more cameras are arranged in a grid such that the cameras view the same side of a person such that the visible surface of that side of a person whose dimensions are the largest dimensions for which the kiosk is designed is captured such that each part of the visible surface of that side of the person is present in one or more of the images which are neighbouring and partially overlapped.

6. A kiosk apparatus according to claim 3, characterised in that two cameras are positioned such that the first camera covers one side of the person and that the second camera is at approximately 90 degrees to the first camera so as to cover an orthogonal side of the person such that for each camera the visible surface of one side of a person whose dimensions are the largest dimensions for which the kiosk is designed is captured on a single image.

7. A kiosk apparatus according to claim 3, characterised in that two sets of two or more cameras are positioned such that the first set of cameras covers one side of the person and that the second set of cameras is at approximately 90 degrees to the first set of cameras so as to cover an orthogonal side of the person such that each set of cameras view the same side of a person such that the visible surface of that side of a person whose dimensions are the largest dimensions for

which the kiosk is designed is captured such that each part of the visible surface of that side of the person is present in one or more of the images from that set of cameras which are neighbouring and partially overlapped.

8. A kiosk apparatus according to any one of claims 2 to 8, characterised in that the illumination is white light.

9. A kiosk apparatus according to any one of claims 2 to 8, characterised in that the illumination is structured light.

10. A kiosk apparatus according to claim 9, characterised in that the structured light is laser stripe.

11. A kiosk apparatus according to any one of claims 2 to 10, characterised in that the illumination is illuminated and each camera takes an image approximately synchronously.

12. A kiosk apparatus according to any one of claims 2 to 11, characterised in that the background as seen by any camera is monochromatic.

13. A kiosk apparatus according to any one of claims 2 to 11, characterised in that the background as seen by any camera is made of a material that is seen by the camera as having a dominant component of a single wavelength when illuminated directly by an illuminating means of the same wavelength.

14. A kiosk apparatus according to claim 13, characterised in that the background material has a retroreflective component.

15. A kiosk apparatus according to any one of claims 13 to 14, characterised in that the illumination means comprises at least two sets of lights each set of lights comprising one or more lights of the same wavelength and each set of lights having a significantly different characteristic wavelength from all the other sets of lights.

16. A kiosk apparatus according to any one of claims 2 to 15, further comprising another camera and position changing means for the camera such that the position of the camera may be changed by the user to frame the user's face when the user is in the correct pose.

17. A kiosk apparatus according to any one of claims 2 to 16, characterised in that the data capture means includes means for laser stripe 3D scanning the person.

18. A kiosk apparatus according to any one of claims 2 to 17, further comprising a display means.
19. A kiosk apparatus according to any one of claims 2 to 18, characterised in that the kiosk is self-contained.
20. A kiosk apparatus according to any one of claims 2 to 19, characterised in that the kiosk is of rugged construction.
21. A kiosk apparatus according to any one of claims 2 to 20, characterised in that the kiosk is of compact design.
22. A kiosk apparatus according to any one of claims 2 to 21, further comprising a weighing means.
23. A kiosk apparatus according to any one of claims 2 to 22, further comprising a network communication means.
24. A kiosk apparatus according to claim 23, characterised in that the kiosk is connected to the internet.
25. A kiosk apparatus according to claim 24, characterised in that there is a server on the internet to which the kiosk transmits the Avatar.
26. A kiosk apparatus according to any one of claims 2 to 25, further comprising:
- a means of recording the Avatar onto a computer readable medium;
  - a means of presenting the computer readable medium to the user.
27. A kiosk apparatus according to any one of claims 2 to 26, further comprising: means for accepting payment.
28. A kiosk apparatus according to claim 27, characterised in that the means for accepting payment is a credit card reading means and a telecommunication connection to a credit card verification means.
29. A kiosk apparatus according to claim 27, characterised in that the means for accepting payment is a means for accepting coins or banknotes.
30. A kiosk apparatus according to any one of claims 2 to 29, further comprising a manual data input means.

31. A kiosk apparatus according to any one of claims 2 to 30, further comprising a sound projection means.

32. A kiosk apparatus according to any one of claims 2 to 31, further comprising graphical guidance means to guide the user in adopting the correct pose or series of poses.

33. A kiosk apparatus according to any one of claims 2 to 32, further comprising advertising means.

34. A kiosk apparatus according to any one of claims 2 to 33, further comprising printing means.

35. A kiosk apparatus according to any one of claims 2 to 34, further comprising sound recording means.

36. A kiosk apparatus according to any one of claims 17 to 35, characterised in that the captured data comprises 3D laser stripe scanned data and images and that the means for generating the Avatar from the captured data comprises an implicit surface avatar generating means.

37. A kiosk apparatus according to claim 36, characterised in that the implicit surface avatar generating means comprises:

triangular sub-mesh creating means;

implicit surface generation means;

hole filling means;

polygon reduction means;

mesh smoothing means;

texture mapping means.

38. A kiosk apparatus according to any one of claims 2 to 35, characterised in that the captured data is images and the means for generating the Avatar from the captured data comprises a silhouette avatar generating means.

39. A kiosk apparatus according to claim 36, characterised in that the silhouette avatar generating means comprises:

a means for extracting silhouettes and landmarks from the images;

a standard Avatar 3D model;

silhouettes and landmarks of the standard Avatar 3D model;

a means for 2D to 2D mapping the silhouettes and landmarks of the standard Avatar 3D model to the silhouettes and landmarks from the captured data;

a means of using the mapped silhouettes and landmarks to deform the standard Avatar 3D model into an untextured Avatar 3D model of the user;

a means of using the mapped silhouettes and landmarks to map the image textures onto the untextured Avatar 3D model of the user to produce a textured Avatar 3D model of the user.

40. A kiosk apparatus according to any one of claims 2 to 39, characterised in that the Avatar is transmitted in a standard format.

41. A kiosk apparatus according to claim 40, characterised in that the standard format is proprietary.

42. A kiosk apparatus according to any one of claims 2 to 41, further comprising means for generating an estimated naked, textured 3D Avatar of the person from data captured in which the person is wearing clothes.

43. A method for generating an Avatar of a person in a kiosk containing a programmable processing apparatus, comprising the following steps:

capturing data on the surface of a person situated inside the kiosk;

processing the data to generate an Avatar;

transmitting the Avatar to the user.

44. A method according to claim 43, characterised in that the kiosk is situated in a public place.

45. A method according to claims 43 and 44, characterised in that one or more Avatars may be generated and one or more Avatars may be transmitted to the user.

46. A method according to any one of claims 43 to 45, characterised in that the Avatar is transmitted to the user by means of supplying the Avatar on a computer readable medium.
47. A method according to claim 46, characterised in that the computer readable medium is a diskette.
48. A method according to claim 46, characterised in that the computer readable medium is a CD-ROM.
49. A method according to any one of claims 43 to 45, characterised in that the Avatar is transmitted to the user by means of transmitting the Avatar to a processing apparatus designated by a user over a network.
50. A method according to claim 49, characterised in that the network is the Internet.
51. A method according to any one of claims 43 to 45, characterised in that the Avatar is transmitted to the user by means of transmitting the Avatar to a server from which the user can thereafter retrieve it over a network.
52. A method according to any one of claims 43 to 51, characterised in that the data on the surface of a person is captured by one or more cameras.
53. A method according to claim 52, characterised in that the person is illuminated whilst the data is captured.
54. A method according to claim 53, characterised in that the illumination is white light.
55. A method according to claim 53, characterised in that the illumination is structured light.
56. A method according to any one of claims 43 to 55, characterised in that the data on the surface of a person is captured with the person standing in one designated pose.
57. A method according to any one of claims 43 to 55, characterised in that the person stands in a series of two or more designated poses and that data on the surface of the person is captured for each pose.
58. A method according to claim 57, characterised in that the person stands in a series of four designated poses.

59. A method according to any one of claims 56 to 58, characterised in that for each pose the person is illuminated sequentially by sets of lights of different characteristic wavelengths and data is captured for each set of lights.

60. A method according to any one of claims 43 to 59, characterised in that the method includes a further step in which the user makes a payment at any time prior to receiving the Avatar.

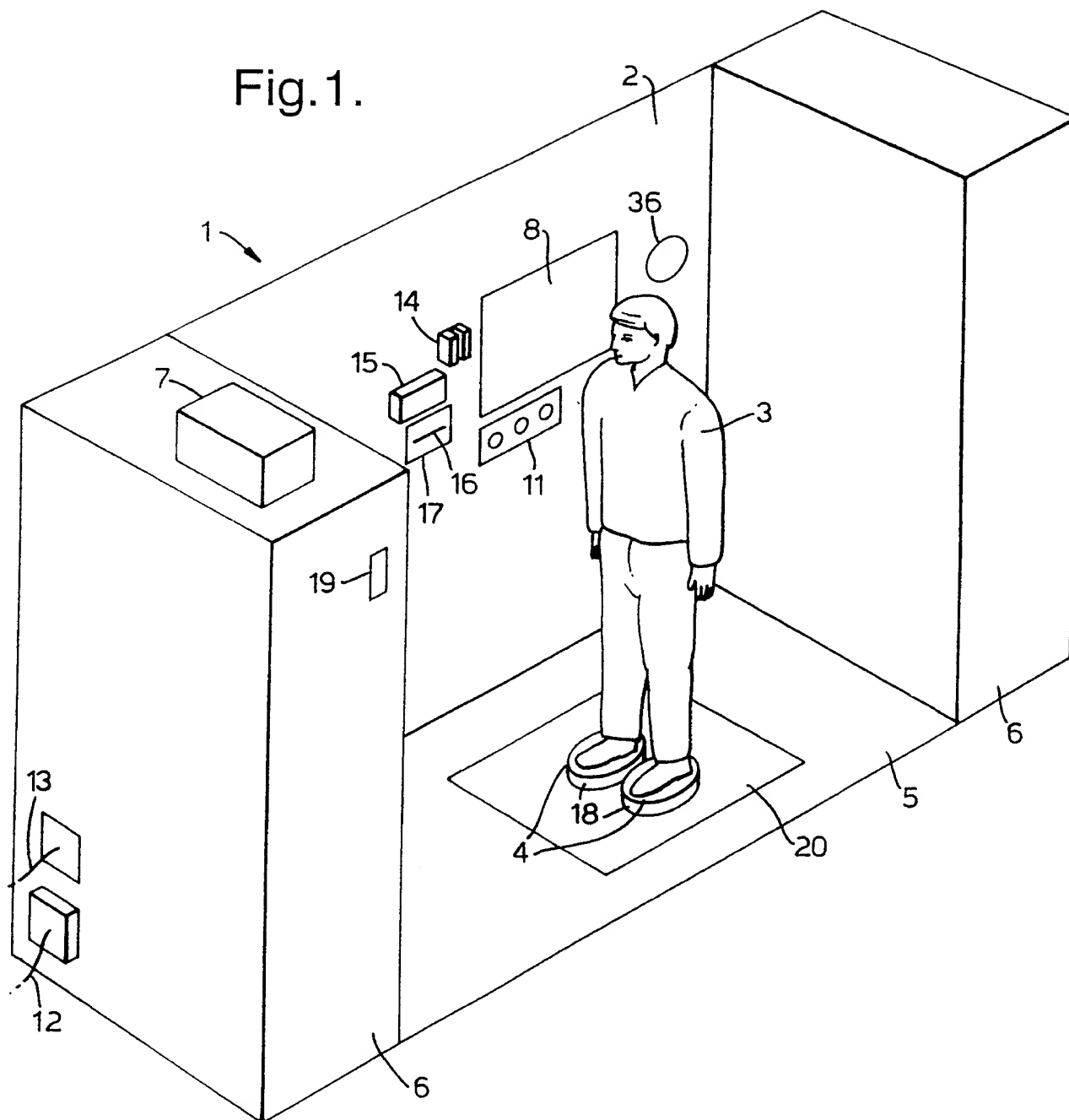
61. A method according to any one of claims 43 to 60, characterised in that the kiosk is operable by the person whose avatar is to be generated without assistance from any other person.

62. A method according to any one of claims 43 to 61, characterised in that a further step in the method generates an accurately estimated naked Avatar.



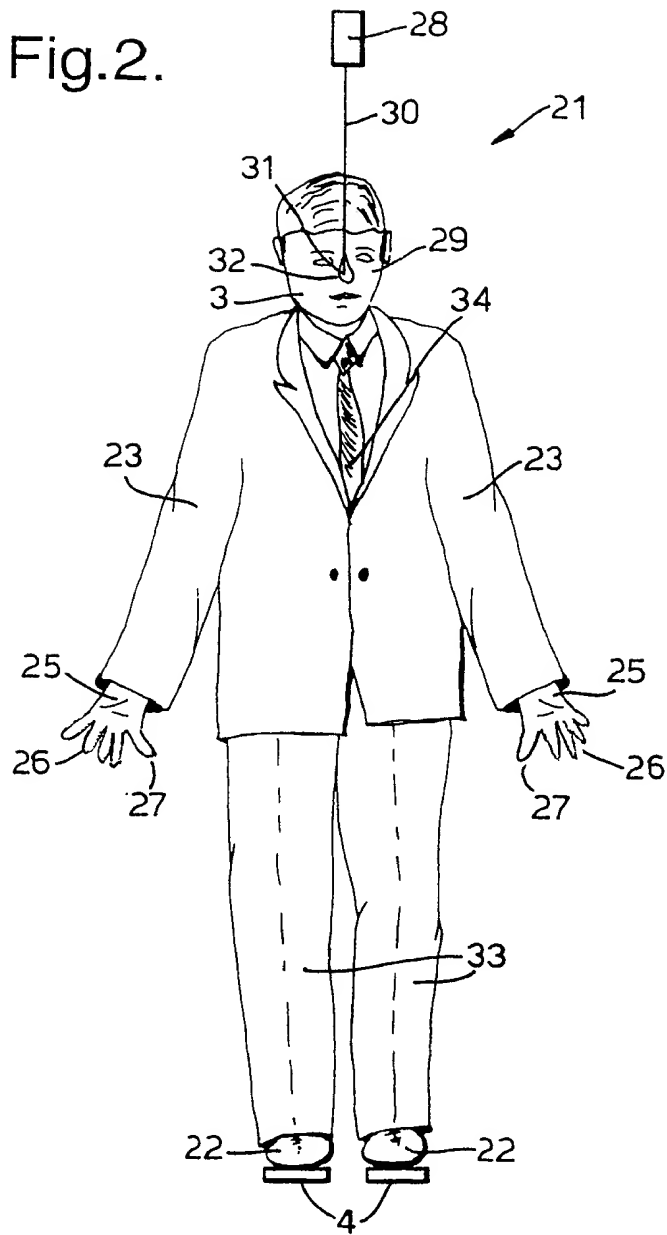
1/9

Fig.1.



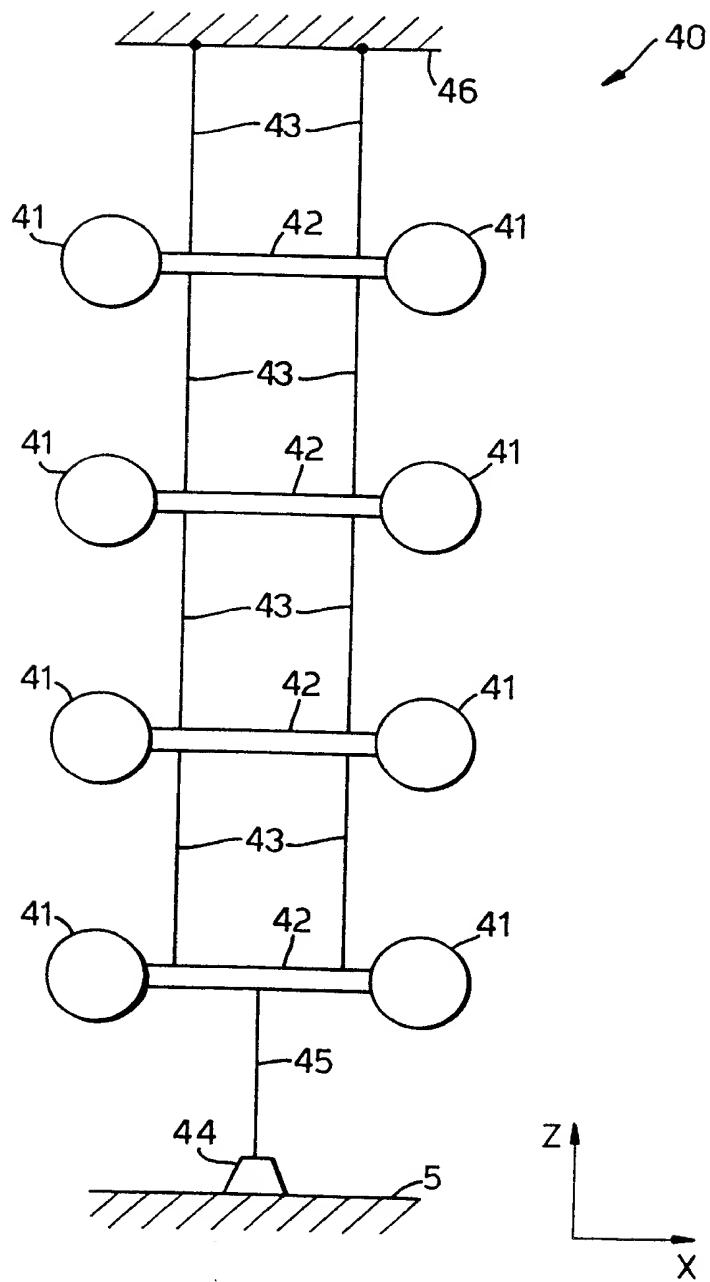
2 / 9

Fig.2.



3/9

Fig.3.



4/9

Fig.4.

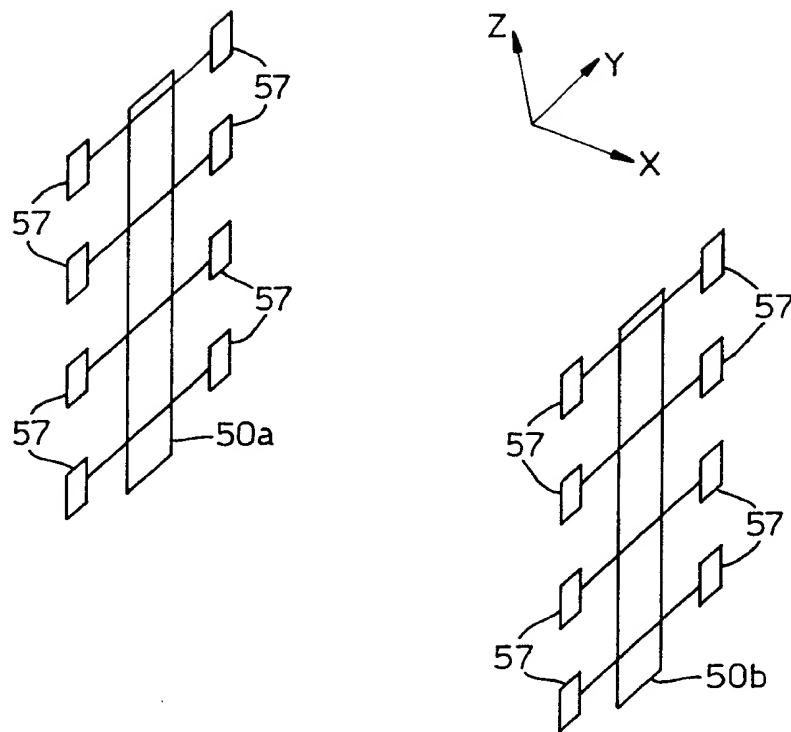
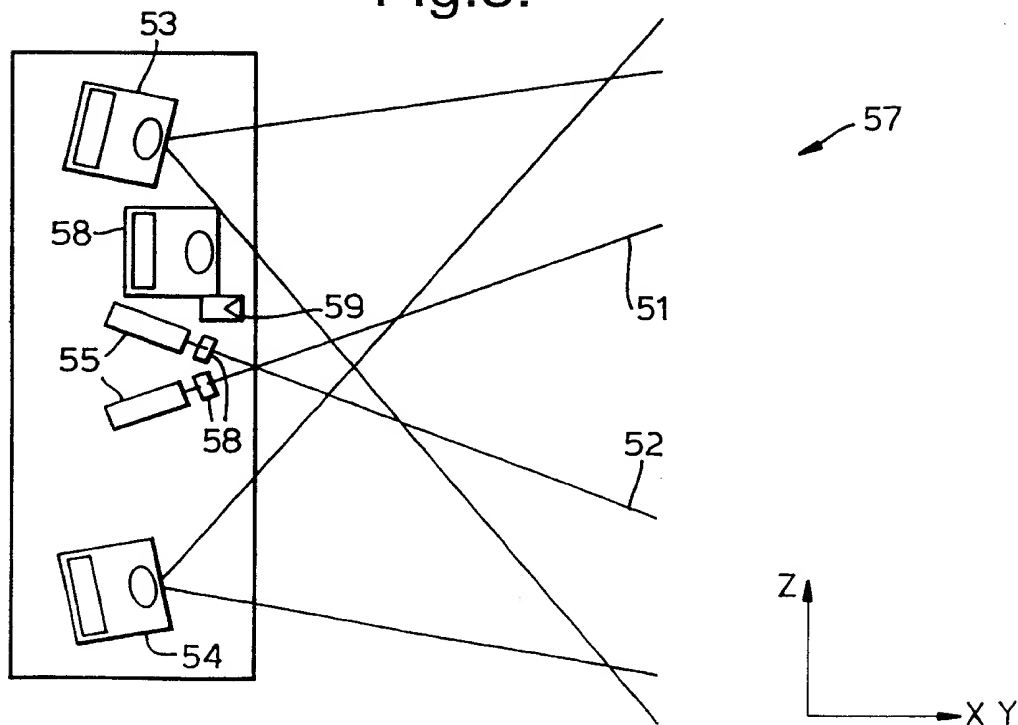


Fig.5.



5/9

Fig.6.

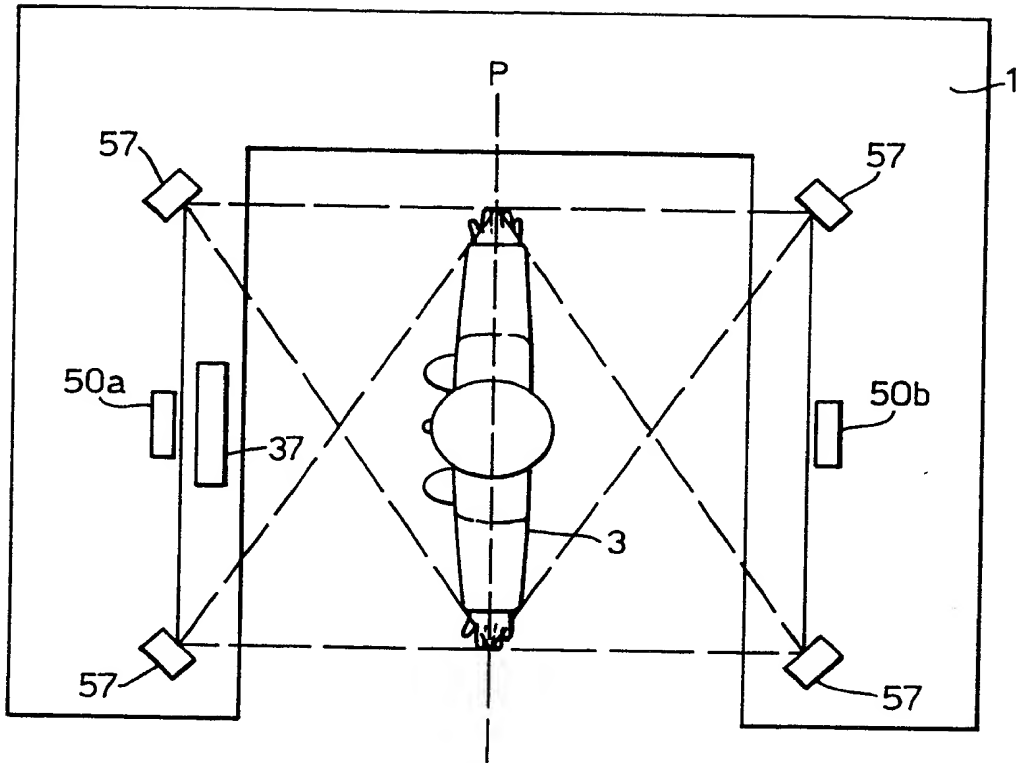
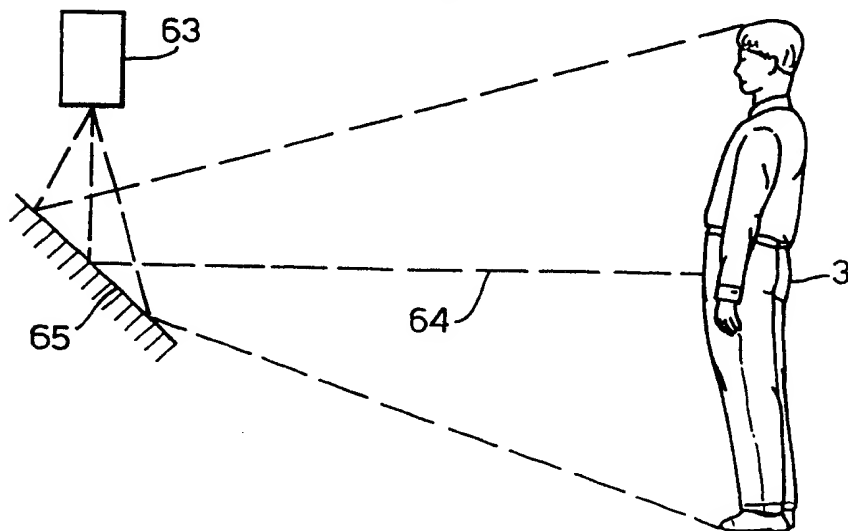
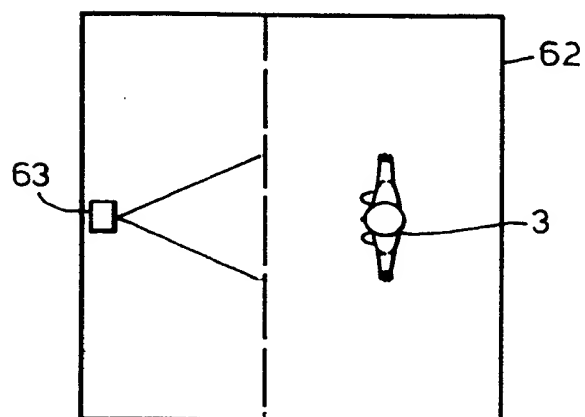
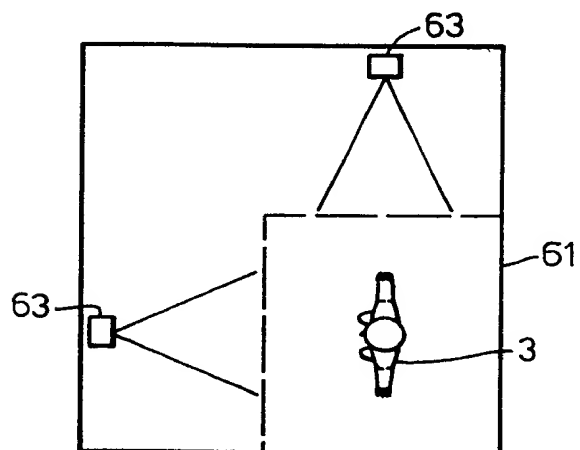
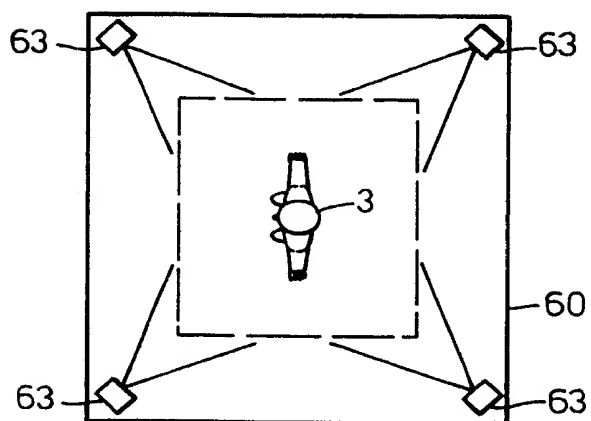


Fig.8.



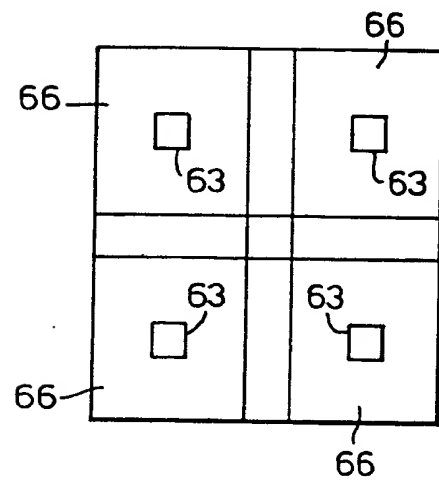
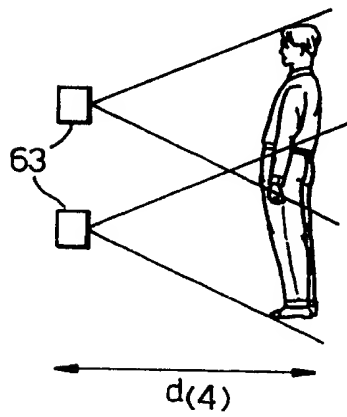
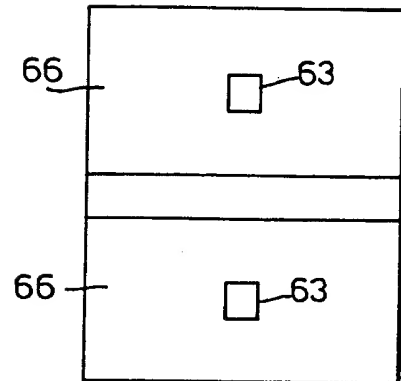
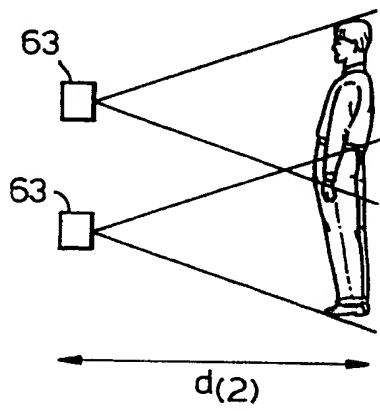
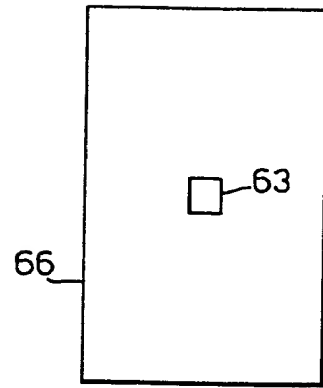
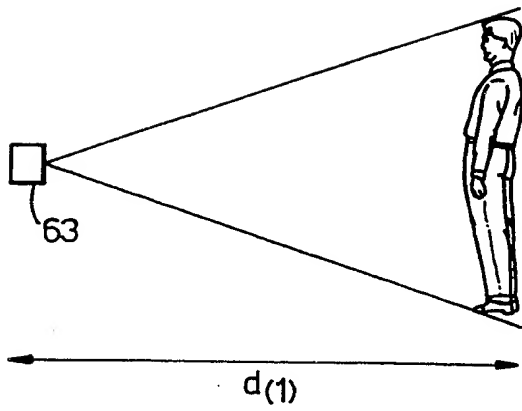
6/9

Fig.7.



7/9

Fig.9.



8/9

Fig.10.

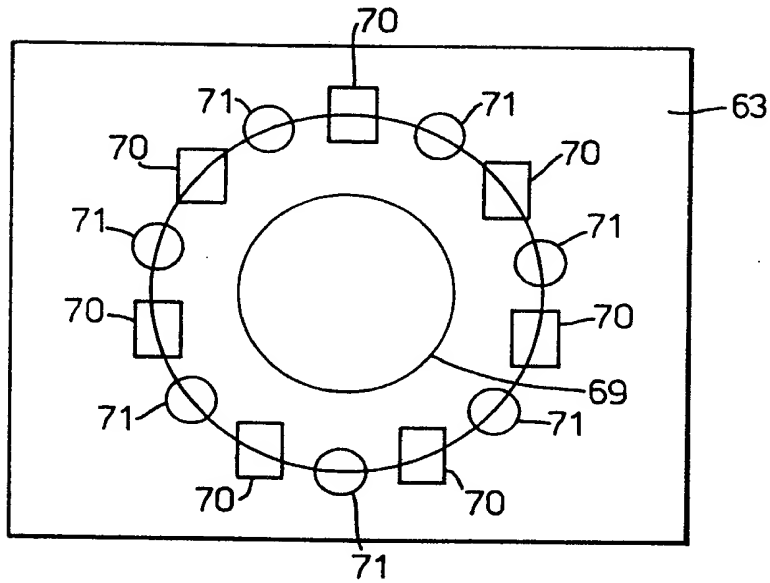
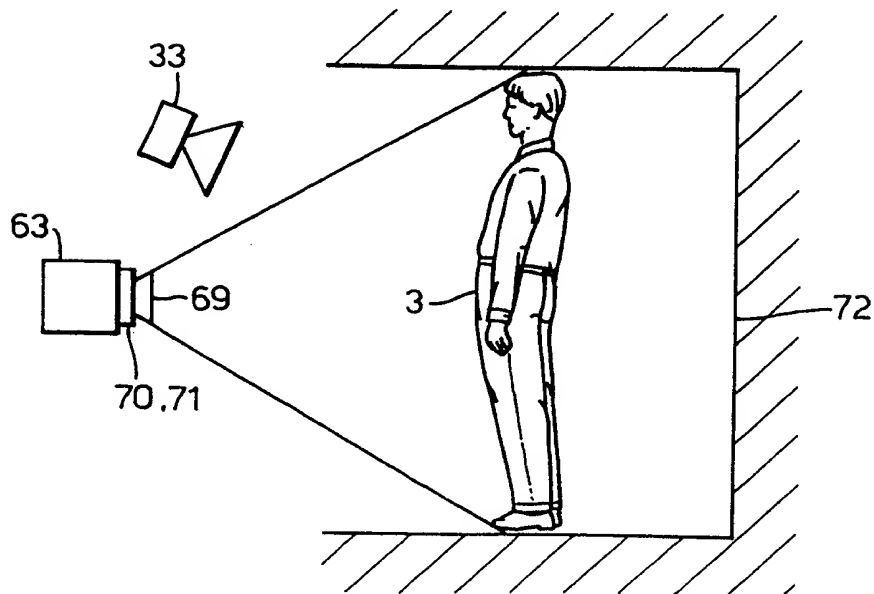


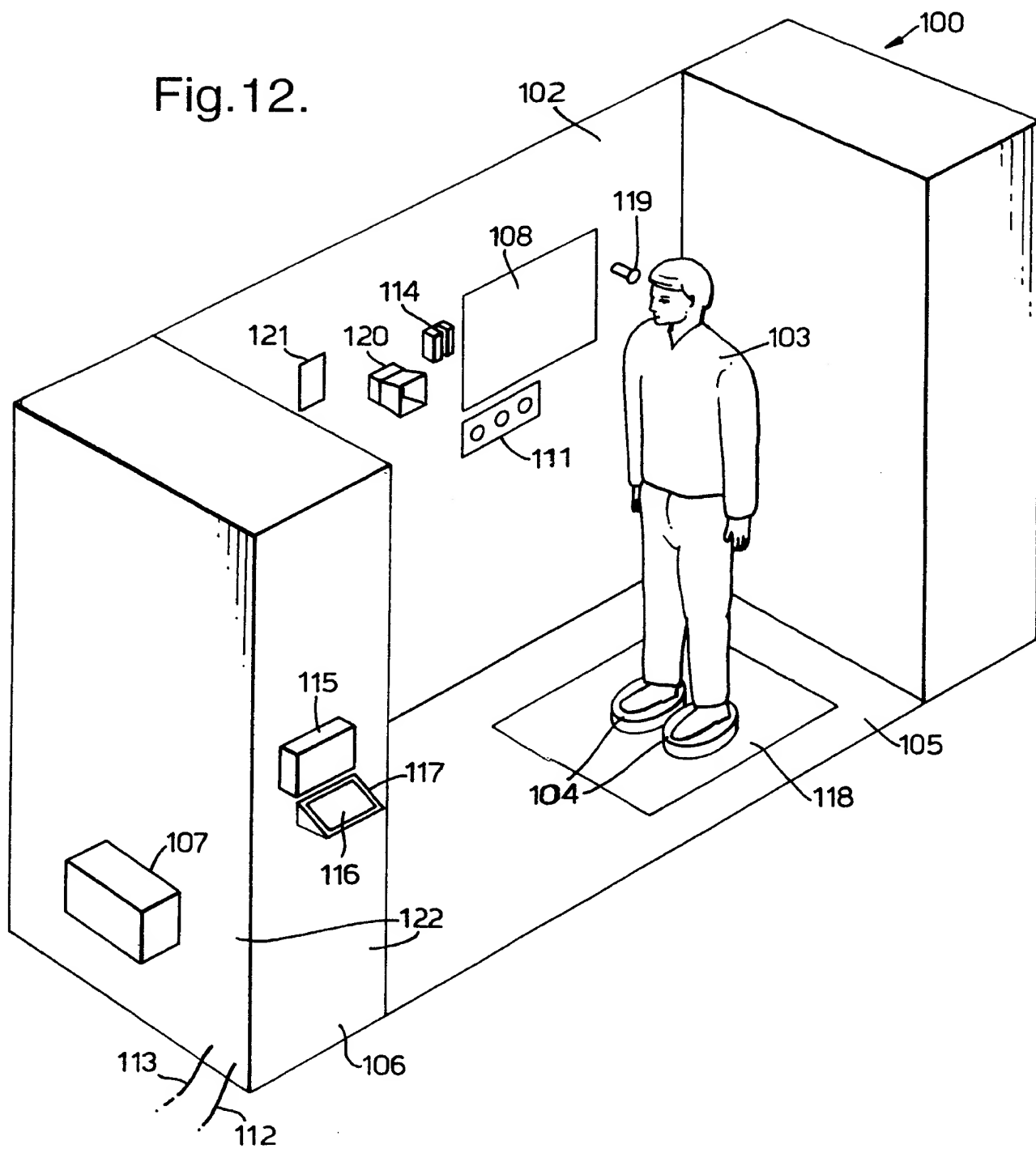
Fig.11.





9/9

Fig.12.



# INTERNATIONAL SEARCH REPORT

International Application No

PCT/GB 97/03515

**A. CLASSIFICATION OF SUBJECT MATTER**  
IPC 6 H04N1/00

According to International Patent Classification(IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	FR 2 724 801 A (KIS) 22 March 1996  see page 11 - page 13 ---	1-4, 43-46, 49
X	"Aiming a camera to take a self portrait" IBM TECHNICAL DISCLOSURE BULLETIN VOL. 35, N°3, PAGES 75-77, US, vol. 35, no. 3, August 1992, pages 75-77, XP000326178 see the whole document ---  -/--	1-4, 43-46, 49

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

13 May 1998

Date of mailing of the international search report

27/05/1998

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Chateau, J-P

## INTERNATIONAL SEARCH REPORT

Inter. Patent Application No.

PCT/GB 97/03515

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	LEE Y ET AL: "Constructing physics-based facial models of individuals" PROCEEDINGS GRAPHICS INTERFACE '93, TORONTO, ONT., CANADA, 19-21 MAY 1993, ISBN 0-9695338-2-9, 1993, TORONTO, ONT., CANADA, CANADIAN INF. PROCESS. SOC, CANADA, pages 1-8, XP002064612 see figure 4	37
A	----- J. ROSS STRENTSTROM ET AL.: "constructing object models from multiple images" INTERNATIONAL JOURNAL OF COMPUTER VISION., vol. 9, no. 3, 1992, NORWELL US, pages 185-212, XP000328501 see page 1, left-hand column, paragraph 1 - right-hand column, paragraph 1.1 -----	1

# INTERNATIONAL SEARCH REPORT

information on patent family members

International Application No

PCT/GB 97/03515

Patent document  
cited in search report

Publication  
date

Patent family  
member(s)

Publication  
date

FR 2724801 A

22-03-96

NONE